

# Iterative Merging Algorithm for Cooperative Data Exchange

Ni Ding\*, Rodney A. Kennedy\* and Parastoo Sadeghi\*

\*Research School of Engineering, College of Engineering and Computer Science, the Australian National University (ANU), Canberra, ACT 2601

Email: {ni.ding, rodney.kennedy, parastoo.sadeghi}@anu.edu.au

**Abstract**—We consider the problem of finding the minimum sum-rate strategy in cooperative data exchange systems that do not allow packet-splitting (NPS-CDE). In an NPS-CDE system, there are a number of geographically close cooperative clients who send packets to help the others recover a packet set. A minimum sum-rate strategy is the strategy that achieves universal recovery (the situation when all the clients recover the whole packet set) with the the minimal sum-rate (the total number of transmissions). We propose an iterative merging (IM) algorithm that recursively merges client sets based on a lower estimate of the minimum sum-rate and updates to the value of the minimum sum-rate. We also show that a minimum sum-rate strategy can be learned by allocating rates for the local recovery in each merged client set in the IM algorithm. We run an experiment to show that the complexity of the IM algorithm is lower than that of the existing deterministic algorithm when the number of clients is lower than 94.

## I. INTRODUCTION

Due to the growing amount of data exchange over wireless networks and increasing number of mobile clients, the base-station-to-peer (B2P) links are severely overloaded. It is called the ‘last mile’ bottleneck problem in wireless transmissions. Cooperative peer-to-peer (P2P) communications is proposed for solving this problem. The idea is to allow mobile clients to exchange information with each other through P2P links instead of solely relying on the B2P transmissions. If the clients are geographically close to each other, the P2P transmissions could be more reliable and faster than B2P ones.

Consider the situation when a base station wants to deliver a set of packets to a group of clients. Due to the fading effects of wireless channels, after broadcast via B2P links, there may still exist some clients that do not obtain all the packets. However, the clients’ knowledge of the packet set may be complementary to each other. Therefore, instead of relying on retransmissions from the base station, the clients can broadcast linear packet combinations of the packets they know via P2P links so as to help the others recover the missing packets. We call this kind of transmission method cooperative data exchange (CDE) and the corresponding system CDE system.

Let the *universal recovery* be the situation that all clients obtain the entire packet set and the *sum-rate* be the total number of linear combinations sent by all clients. In CDE systems, the most commonly addressed problem is to find the minimum-sum rate strategy, the transmission scheme that achieves universal recovery and has the minimum sum-rate.

This problem was introduced in [1]. Randomized and deterministic algorithms for solving this problem have been proposed in [2], [3] and [4], [5], respectively. The idea of the randomized algorithms in [2], [3] is to choose a client with the maximal or non-minimal rank of the received encoding vectors and let him/her transmit once by using random coefficients from a large Galois field. But, these randomized algorithms repetitively call the rank function, the complexity of which grows with both the number of clients and the number of packets. On the other hand, the authors in [4], [5] propose deterministic algorithms where the complexity only grows with the number of clients. But, we will show in this paper that the divide-and-conquer (DV) algorithm proposed in [4] can not be applied to CDE systems that do not allow packet-splitting (NPS-CDE). Although the deterministic algorithm in [5] can solve NPS-CDE problems, it relies on the submodular function minimization (SFM) algorithm, and the complexity of SFM algorithms is not low.<sup>1</sup>

In this paper, we first use a counter example to show that the DV algorithm in [4] can not solve the NPS-CDE problems. We then propose an iterative merging (IM) algorithm, a deterministic algorithm, for finding the minimum sum-rate and corresponding strategy in NPS-CDE systems. The IM algorithm starts with an initial lower estimate of the minimum sum-rate. It recursively merges the clients that require the least number of transmissions for both the local recovery and the recovery of the collectively missing packets.<sup>2</sup> The IM algorithm updates the estimate of minimum sum-rate whenever it finds that the universal recovery is not achievable. We prove that the minimum sum-rate can be found by starting the IM algorithm. We also show that a minimum sum-rate strategy can be determined by allocating transmission rates for the local recovery in each merged client sets in IM algorithm. We run an experiment to show that the complexity of the IM algorithm is lower than that of the deterministic algorithm proposed in [5] when the number of clients is lower than 94.

<sup>1</sup>There are many algorithms proposed for solving SFM problems. To our knowledge, the algorithm proposed in [6] has the lowest complexity  $O(K^5 \cdot \gamma + K^6)$ , where  $K$  is the number of clients, and  $\gamma$  is the complexity of evaluating a submodular function.

<sup>2</sup>Local recovery means the merged clients exchange whatever missing in the packet set that they collectively know so that they share the same common knowledge and can be treated as a single entity.

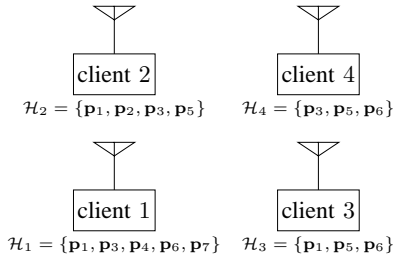


Fig. 1. An example of CDE system where there are four clients that want to obtain eight packets.  $\mathcal{H}_j$  is the has-set of client  $j$ .

## II. SYSTEM MODEL AND PROBLEM STATEMENT

Let  $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_L\}$  be the packet set containing  $L$  linearly independent packets. Each packet  $\mathbf{p}_i$  belongs to the finite field  $\mathbb{F}_q$ . The system contains  $K$  geographically close clients. Let  $\mathcal{K} = \{1, \dots, K\}$  be the client set. Each client  $j \in \mathcal{K}$  initially obtains  $\mathcal{H}_j \subset \mathcal{P}$ . Here,  $\mathcal{H}_j$  is called the *has-set* of client  $j$ . The clients are assumed to collectively know the packet set, i.e.,  $\cup_{j \in \mathcal{K}} \mathcal{H}_j = \mathcal{P}$ . The P2P wireless links between clients are error-free, i.e., information broadcast by any client can be heard losslessly by all other clients. The clients broadcast linear combinations of the packets in their has-sets in order to help each other recover  $\mathcal{P}$ . For example, in the CDE system in Fig. 1, client 1 broadcasting  $\mathbf{p}_1 + \mathbf{p}_6$  helps client 2 and client 4 recover  $\mathbf{p}_6$  and  $\mathbf{p}_1$ , respectively. Assume packet-splitting is not allowed. Let  $\mathbf{r} = (r_1, \dots, r_K)$  be a transmission strategy with  $r_j \in \mathbb{N}_0$  being the total number of linear combinations transmitted by client  $j$ . We call  $\sum_{j \in \mathcal{K}} r_j$  the *sum-rate* of strategy  $\mathbf{r}$ . Let the *universal recovery* be the situation that all clients in  $\mathcal{K}$  obtains the entire packet set  $\mathcal{P}$ . The problem is to find a *minimum sum-rate transmission strategy*, a strategy that has the minimum sum-rate among all strategies that achieve universal recovery.

## III. MINIMUM SUM-RATE STRATEGY

In this section, we first clarify the notations, or definitions, that used in this paper and then discuss how to determine the minimum sum-rate and a minimum sum-rate strategy.

Denote  $\mathcal{W}_{\mathcal{K}}$  a *partition* of the client set  $\mathcal{K}$ .<sup>3</sup> Let  $|\mathcal{W}_{\mathcal{K}}|$  be the cardinality of  $\mathcal{W}_{\mathcal{K}}$ . We call  $\mathcal{W}_{\mathcal{K}}$  a  $|\mathcal{W}_{\mathcal{K}}|$ -partition of  $\mathcal{K}$ . For example,  $\mathcal{W}_{\mathcal{K}} = \{\{1, 2\}, \{3\}, \{4\}\}$  is a 3-partition of client set  $\mathcal{K} = \{1, 2, 3, 4\}$ . Let  $\mathcal{Y} \subseteq \mathcal{W}_{\mathcal{K}}$ . We call  $\mathcal{Y}$  the  $k$ -subset of partition  $\mathcal{W}_{\mathcal{K}}$  if  $|\mathcal{Y}| = k$ , e.g.,  $\{\{2\}, \{3, 4\}\}$  is a 2-subset of partition  $\mathcal{W}_{\mathcal{K}} = \{\{1\}, \{2\}, \{3, 4\}\}$ . Let  $\mathcal{Y} = \cup_{\mathcal{X} \in \mathcal{Y}} \mathcal{X}$ , e.g., if  $\mathcal{Y} = \{\{2\}, \{3, 4\}\}$ ,  $\mathcal{Y} = \{2, 3, 4\}$ .

For  $\mathcal{S} \subseteq \mathcal{K}$ , denote  $\mathbf{r}_{\mathcal{S}} = \sum_{j \in \mathcal{S}} r_j$  and  $\mathcal{H}_{\mathcal{S}} = \cup_{j \in \mathcal{S}} \mathcal{H}_j$ . We define the *local recovery* in  $\mathcal{S}$  as the situation such that all clients  $j \in \mathcal{S}$  obtain  $\mathcal{H}_{\mathcal{S}}$ . For example, in Fig. 1, for  $\mathcal{S} = \{3, 4\}$ , the problem of local recovery is how to let both client 3 and client 4 obtain the packet set  $\mathcal{H}_{\{3, 4\}} = \{\mathbf{p}_1, \mathbf{p}_3, \mathbf{p}_5, \mathbf{p}_6\}$ . The minimum sum-rate  $\alpha_{\mathcal{S}}^*$  for the local recovery in  $\mathcal{S}$  is

determined by [7]

$$\alpha_{\mathcal{S}}^* = \max \left\{ \left\lceil \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}} \frac{|\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|}{|\mathcal{W}_{\mathcal{S}}| - 1} \right\rceil : \mathcal{W}_{\mathcal{S}} \text{ is a partition of } \mathcal{S} \text{ that satisfies } 2 \leq |\mathcal{W}_{\mathcal{S}}| \leq |\mathcal{S}| \right\}. \quad (1)$$

(1) is based on the condition that  $\mathbf{r}_{\mathcal{X}} \geq |\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|$  must be satisfied for all  $\mathcal{X} \subset \mathcal{S}$  for the local recovery in  $\mathcal{S}$  [5].<sup>4</sup> An equivalent interpretation of (1) is that  $\alpha_{\mathcal{S}}^*$  is the minimum integer that satisfies

$$\alpha_{\mathcal{S}}^* \leq \min \left\{ \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}} (\alpha_{\mathcal{S}}^* - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|) : \mathcal{W}_{\mathcal{S}} \text{ is a partition of } \mathcal{S} \text{ that satisfies } 2 \leq |\mathcal{W}_{\mathcal{S}}| \leq |\mathcal{S}| \right\}. \quad (2)$$

*Proposition 3.1:* In an NPS-CDE system, let  $\mathcal{S} \subseteq \mathcal{K}$ ,  $\Delta\alpha_{\mathcal{S}} = \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}^*} (\alpha_{\mathcal{S}}^* - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|) - \alpha_{\mathcal{S}}^*$  and  $\mathcal{X}'$  be any subset in  $\mathcal{W}_{\mathcal{S}}^*$ . A strategy that achieves local recovery in  $\mathcal{S}$  satisfies

$$\begin{aligned} \mathbf{r}_{\mathcal{X}} &= \alpha_{\mathcal{S}}^* - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|, \forall \mathcal{X} \in \mathcal{W}_{\mathcal{S}}^* \setminus \mathcal{X}', \\ \mathbf{r}_{\mathcal{X}'} &= \alpha_{\mathcal{S}}^* - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}'}| - \Delta\alpha_{\mathcal{S}}. \end{aligned}$$

*Proof:* The constraint conditions  $\{\mathbf{r}_{\mathcal{X}} \geq |\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|, \forall \mathcal{X} \subset \mathcal{K}, \mathbf{r}_{\mathcal{K}} = \alpha_{\mathcal{S}}^*\}$  are equivalent to  $\{\mathbf{r}_{\mathcal{X}} \leq \alpha_{\mathcal{S}}^* - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|, \forall \mathcal{X} \subset \mathcal{K}, \mathbf{r}_{\mathcal{K}} = \alpha_{\mathcal{S}}^*\}$ . Among all these constraints, the tightest ones are  $\{\mathbf{r}_{\mathcal{X}} \leq \alpha_{\mathcal{S}}^* - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|, \forall \mathcal{X} \subset \mathcal{W}_{\mathcal{S}}^*, \mathbf{r}_{\mathcal{K}} = \alpha_{\mathcal{S}}^*\}$ , and the excessive rate  $\Delta\alpha_{\mathcal{S}}$  can be reduced from any client set in  $\mathcal{W}_{\mathcal{S}}^*$ . So, proposition holds. ■

It can be seen that the universal recovery is also the local recovery in  $\mathcal{K}$ , i.e., (1) and Proposition 3.1 can be applied for universal recovery by letting  $\mathcal{S} = \mathcal{K}$ . We call  $\mathcal{W}_{\mathcal{S}}^*$  the minimum sum-rate partition for the local recovery in  $\mathcal{S}$ . Let  $\mathcal{W}_{\mathcal{S}}^{\circ}$  be the maximizer of (1). It should be noted that  $\mathcal{W}_{\mathcal{S}}^{\circ}$  and  $\mathcal{W}_{\mathcal{S}}^*$  are not necessarily equal and  $\mathcal{W}_{\mathcal{S}}^{\circ}$  can not be used for determining the minimum sum-rate for the local recovery in  $\mathcal{S}$ . See the Example 4.1 in the next section.

## IV. ERRORS IN DIVIDE-AND-CONQUER ALGORITHM

The authors in [4] proposed a divide-and-conquer (DV) algorithm for finding the minimum sum-rate strategy in NPS-CDE systems.<sup>5</sup> In Theorem 1 in [4], it states that the minimum sum-rate for the local recovery in  $\mathcal{S}$  is given by

$$\alpha_{\mathcal{S}}^{\circ} = \max \left\{ \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}} \frac{|\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|}{|\mathcal{W}_{\mathcal{S}}| - 1} : \mathcal{W}_{\mathcal{S}} \text{ is a partition of } \mathcal{S} \text{ that satisfies } 2 \leq |\mathcal{W}_{\mathcal{S}}| \leq |\mathcal{S}| \right\}. \quad (3)$$

In Lemma 1 in [4], it states that the minimum sum-rate transmission strategy for the local recovery in  $\mathcal{S}$  can be determined by

$$\sum_{j \in \mathcal{X}} r_j = \alpha_{\mathcal{S}}^{\circ} - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|, \forall \mathcal{X} \in \mathcal{W}_{\mathcal{S}}^{\circ}, \quad (4)$$

where  $\mathcal{W}_{\mathcal{S}}^{\circ}$  is the maximizer of (3). However, in most of the cases,  $\alpha_{\mathcal{S}}^{\circ}$  is not an integer and Lemma 1 in [4] is not

<sup>3</sup>A partition  $\mathcal{W}_{\mathcal{K}}$  satisfies  $\emptyset \neq \mathcal{X} \subseteq \mathcal{K}$ ,  $\mathcal{X} \cap \mathcal{X}' = \emptyset$  and  $\cup_{\mathcal{X} \in \mathcal{W}_{\mathcal{K}}} \mathcal{X} = \mathcal{K}$  for all  $\mathcal{X}, \mathcal{X}' \in \mathcal{W}_{\mathcal{K}}$ .

<sup>4</sup>A brief proof of (1) is given in Appendix A.

<sup>5</sup>See details of DV algorithm in Appendix B.

correct. For NPS-CDE systems, the minimum sum-rate must be an integer since every client must transmit integer number of times. So,  $\alpha_S^\circ$  determined by (3) can not necessarily be the minimum sum-rate. It should be round up to a closest integer as in (1).<sup>6</sup> One may suggest replacing  $\alpha_S^\circ$  by  $\alpha_S^*$  in (4). However, if so, Lemma 1 in [4] does not hold. See the example below.

*Example 4.1:* Consider a CDE system in Fig. 1. For the universal recovery, the maximum and the maximizer of (3) are  $\alpha_S^\circ = 13/3$  and  $\mathcal{W}_S^\circ = \{\{1\}, \{2\}, \{3\}, \{4\}\}$ , respectively. The corresponding transmission strategy is  $\mathbf{r} = (7/3, 4/3, 1/3, 1/3)$  by using (4). This strategy can not be implemented in an NPS-CDE system. Therefore, Lemma 1 in [4] is not correct in this case. In addition, if we use  $\alpha_S^*$  instead of  $\alpha_S^\circ$  in (4), we get  $\mathbf{r} = (3, 2, 1, 1)$ , which achieves universal recovery but has a sum-rate greater than  $\alpha_K^* = 5$ , i.e., it is not a minimum sum-rate strategy. Therefore, Lemma 1 does not hold, either. One may think that reducing 2 transmissions from any client in strategy  $\mathbf{r} = (3, 2, 1, 1)$  will result in a minimum sum-rate transmission strategy. This is also not true. For example, if we reduce the transmission rate of client 1 by 2, we get  $\mathbf{r} = (1, 2, 1, 1)$ . It has sum-rate equals  $\alpha_K^* = 5$ . But, the universal recovery is not achievable since constraint  $r_1 \geq L - |\mathcal{H}_{\{2,3,4\}}| = 2$  is breached. In fact, the correct way is to break  $\mathcal{K}$  into  $\mathcal{W}_S^* = \{\{1, 2, 3\}, \{4\}\}$  and determine the individual rates of the clients in  $\{1, 2, 3\}$  for the local recovery in  $\{1, 2, 3\}$ .<sup>7</sup> We will show that this can be accomplished by IM algorithm in Example 5.1 in the next section.

## V. ITERATIVE MERGING SCHEDULING METHOD

In this section, we propose a greedy scheduling method for the universal recovery in CDE systems. We assume that the clients in CDE system can form coalitions, or groups. A coalition can contain just one client, and each client must appear in no more than one coalition. Any form of coalition in  $\mathcal{K}$  can be represented by a partition  $\mathcal{W}_K$ , and any  $k$ -subset  $\mathcal{Y}$  of  $\mathcal{W}_K$  contains  $k$  coalitions in  $\mathcal{W}_K$ . The idea of this scheduling method is to iteratively merge coalitions and check if condition (2) holds.

Let  $\alpha$  be a lower estimate of  $\alpha_K^*$ , e.g., the lower bound on  $\alpha_K^*$  proposed in [2], [7]. At the beginning, we assume that each client forms one coalition, which can be denoted by a  $K$ -partition  $\mathcal{W}_K = \{\{j\} : j \in \mathcal{K}\}$ . We start an iterative procedure. In each iteration, we perform two steps:

1. If  $\alpha > \sum_{\mathcal{X} \in \mathcal{W}_K} (\alpha - L + |\mathcal{H}_X|)$ , terminate iteration, increase  $\alpha$  by one and start the IM scheduling method (from  $K$ -partition) again; Otherwise, go to step 2.
2. Let  $k \in \{2, \dots, |\mathcal{W}_K|\}$ . We choose  $\mathcal{Y}$  as a  $k$ -subset with the minimum value of  $k$  that satisfies the conditions

$$\sum_{\mathcal{X} \in \mathcal{Y}} \frac{|\mathcal{H}_{\tilde{\mathcal{Y}}}| - |\mathcal{H}_X|}{|\mathcal{Y}| - 1} + L - |\mathcal{H}_{\tilde{\mathcal{Y}}}| < \alpha, \quad (5)$$

<sup>6</sup>In fact,  $\alpha_S^*$  determined by (3) is the minimum sum-rate for CDE systems that allow packet-splitting (PS-CDE). There is a study in [7] shows how to determine minimum sum-rate for both PS-CDE and NPS-CDE systems.

<sup>7</sup>See the explanation in Appendix C.

$$\begin{aligned} & \sum_{\mathcal{X} \in \mathcal{Y}} \frac{|\mathcal{H}_{\tilde{\mathcal{Y}}}| - |\mathcal{H}_X|}{|\mathcal{Y}| - 1} + L - |\mathcal{H}_{\tilde{\mathcal{Y}}}| \\ & \leq \sum_{\mathcal{X} \in \mathcal{Y}'} \frac{|\mathcal{H}_{\tilde{\mathcal{Y}}}| - |\mathcal{H}_X|}{|\mathcal{Y}'| - 1} + L - |\mathcal{H}_{\tilde{\mathcal{Y}}}|, \quad (6) \end{aligned}$$

for all other subsets  $\mathcal{Y}'$  such that  $|\mathcal{Y}| = |\mathcal{Y}'|$ . Achieve local recovery in  $\tilde{\mathcal{Y}}$ , and update  $\mathcal{W}_K$  by merging all coalitions  $\mathcal{X}$  in  $\mathcal{Y}$  into one coalition.

The iteration terminates whenever we find that  $|\mathcal{W}_K| = 2$  or there is no  $k$ -subset  $\mathcal{Y}$  satisfies conditions (5) and (6) in step 2.

Consider step 1. As discussed in Section III, if we find that condition

$$\alpha \leq \sum_{\mathcal{X} \in \mathcal{W}_K} (\alpha - L + |\mathcal{H}_X|) \quad (7)$$

does not hold for some partition  $\mathcal{W}_K$ , it means  $\alpha < \alpha_K^*$  and universal recovery is not possible with the sum-rate  $\alpha$ . Therefore,  $\alpha$  should be increased.

In step 2, the interpretations of the conditions (5) and (6) are as follows. Based on Condition (6),  $\mathcal{Y}$  must be the minimum sum-rate partition for the local recovery of the collectively known packets in  $\tilde{\mathcal{Y}}$ , i.e.,  $\mathcal{Y} = \mathcal{W}_{\tilde{\mathcal{Y}}}^*$ .<sup>8</sup> So,  $\sum_{\mathcal{X} \in \mathcal{Y}} \frac{|\mathcal{H}_{\tilde{\mathcal{Y}}}| - |\mathcal{H}_X|}{|\mathcal{Y}| - 1}$  incurs the minimum sum-rate for the local recovery in  $\tilde{\mathcal{Y}}$ .  $L - |\mathcal{H}_{\tilde{\mathcal{Y}}}|$  is the number of collectively missing packets the recovery of which relies on the transmissions in client set  $\mathcal{K} \setminus \tilde{\mathcal{Y}}$ . If condition (5) is breached, it means that universal recovery with sum-rate  $\alpha$  is not possible if the coalitions in  $\mathcal{Y}$  are merged to form one coalition  $\tilde{\mathcal{Y}}$ . Therefore, it is better for them to work individually than together. Condition (6) means that  $\mathcal{Y}$  require less number of transmissions for the recovery of both collectively known and collectively missing packets than any other  $\mathcal{Y}'$  such that  $|\mathcal{Y}'| = |\mathcal{Y}|$ .

*Example 5.1:* Consider the CDE system in Fig. 1 and assume packet-splitting is not allowed. Let  $\alpha = 5$  and apply IM scheduling method. We get the procedure below.

- Assume that each client works individually at the beginning and initiate  $\mathcal{W}_K = \{\{1\}, \{2\}, \{3\}, \{4\}\}$ . In this case, we have  $\sum_{\mathcal{X} \in \mathcal{W}_K} (\alpha - L + |\mathcal{H}_X|) = 7 > \alpha$ , we continue to step 2 to consider conditions (5) and (6) to determine which coalitions should be merged. It can be shown that there is no 2-subset but one 3-subset  $\{1, 2, 3\}$  satisfies both conditions. Therefore, coalitions  $\{1\}$ ,  $\{2\}$  and  $\{3\}$  should be merged to form one coalition  $\{1, 2, 3\}$ . Consider how to achieve the local recovery in  $\{1, 2, 3\}$ . It can be shown that  $\alpha_{\{1,2,3\}}^* = \lceil \sum_{\mathcal{X} \in \{\{1\}, \{2\}, \{3\}\}} \frac{|\mathcal{H}_{\{1,2,3\}}| - |\mathcal{H}_X|}{|\{\{1\}, \{2\}, \{3\}\}| - 1} \rceil = 5$  and the minimizer of (2) is  $\mathcal{W}_{\{1,2,3\}}^* = \{\{1\}, \{2\}, \{3\}\}$ . In this case,  $\Delta\alpha_{\{1,2,3\}} = \sum_{\mathcal{X} \in \mathcal{W}_{\{1,2,3\}}^*} (\alpha_{\{1,2,3\}}^* - |\mathcal{H}_S| + |\mathcal{H}_X|) - \alpha_{\{1,2,3\}}^* = 1$ . According to Proposition 3.1, we choose to subtract  $\Delta\alpha_{\{1,2,3\}}$  from the rate of client 3 and get rates

$$r_1 = \alpha_{\{1,2,3\}}^* - |\mathcal{H}_{\{1,2,3\}}| + |\mathcal{H}_1| = 3,$$

$$r_2 = \alpha_{\{1,2,3\}}^* - |\mathcal{H}_{\{1,2,3\}}| + |\mathcal{H}_2| = 2,$$

$$r_3 = \alpha_{\{1,2,3\}}^* - |\mathcal{H}_{\{1,2,3\}}| + |\mathcal{H}_3| - \Delta\alpha_{\{1,2,3\}} = 0.$$

<sup>8</sup>We will show that this is the case in the proof of Theorem 5.3 in Section V-A

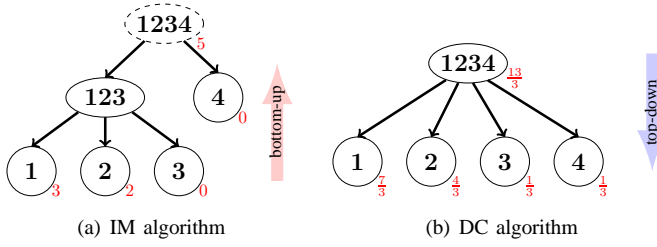


Fig. 2. The merging process results from iterative merging (IM) algorithm and dividing process results from divide-and-conquer (DC) algorithm when they are applied to find the minimum sum-rate strategy in the CDE system in Fig. 1. Note, final merging to one coalition  $\{1, 2, 3, 4\}$  does not happen but is implied in the IM algorithm. In each figure, the minimum sum-rate  $\alpha_K^*$  is shown beside the coalition  $\{1, 2, 3, 4\}$ , and the rates of clients in minimum sum-rate strategy are shown beside singleton coalitions. Note, the strategy determined by DC algorithm can not be implemented in an NPS-CDE system.

- For  $\mathcal{W}_K = \{\{1, 2, 3\}, \{4\}\}$ ,  $\sum_{\mathcal{X} \in \mathcal{W}_K} (\alpha - L + |\mathcal{H}_X|) = 6 > \alpha$ . However, we do not need to determine the merging candidates since the 2-partition will be necessarily merged to form coalition  $\{1, 2, 3, 4\}$ . It is straightforward to see that coalition  $\{1, 2, 3\}$  transmitting 4 times and coalition  $\{4\}$  keeping silent achieve universal recovery. But, there are already 5 transmissions in  $\{1, 2, 3\}$  when achieving local recovery, which means that client 4 has recovered the missing packets by listening to the transmissions for the local recovery in  $\{1, 2, 3\}$ , i.e.,  $\mathbf{r} = (3, 2, 0, 0)$  achieves universal recovery.

We will show that  $\mathbf{r} = (3, 2, 0, 0)$  is a minimum sum-rate strategy and  $\alpha_K^* = 5$  in Theorem 5.3 in the next sections. Note, this procedure also shows a method to determine a minimum strategy: allocate the transmission rates for local recovery in each merged coalition. In Fig. 2, we show the merging and dividing processes incurred by IM and DV algorithms, respectively.

*Example 5.2:* Consider applying the IM scheduling method in the CDE system in Fig. 1 with  $\alpha = 4$ . For  $\mathcal{W}_K = \{\{1\}, \{2\}, \{3\}, \{4\}\}$ ,  $\sum_{\mathcal{X} \in \mathcal{W}_K} (\alpha - L + |\mathcal{H}_X|) = 3 < \alpha$ .  $\alpha$  will be increased to 5 in the first iteration and IM algorithm starts over again where the same procedure as in Example 5.1 is repeated.

#### A. Iterative Merging Algorithm

We describe the IM scheduling method as the IM algorithm in Algorithms 1, 2 and 3. In Algorithm 1,  $v_\alpha$  is defined as

$$v_\alpha(\mathcal{X}) = \alpha - L + |\mathcal{H}_X|, \quad (8)$$

and  $\alpha > \sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$  is equivalent to the breach of condition (7). In Algorithm 2,

$$\xi_\alpha(\mathcal{Y}) = v_\alpha(\tilde{\mathcal{Y}}) - \sum_{\mathcal{X} \in \mathcal{Y}} v_\alpha(\mathcal{X}). \quad (9)$$

$\xi_\alpha(\mathcal{Y}) < 0$  and  $\xi_\alpha(\mathcal{Y}) \leq \xi_\alpha(\mathcal{Y}')$  are equivalent to conditions (5) and (6), respectively.

**Theorem 5.3:** The IM algorithm returns  $\alpha_K^*$  and a minimum sum-rate strategy if the input  $\alpha \leq \alpha_K^*$ .

*Proof:* Consider the Queyranne's algorithm [8]

$$\mathcal{M} := \mathcal{M} \cup \{e\}$$

#### Algorithm 1: Iterative Merge (IM)

---

**input :**  $\alpha$ , a lower bound on  $\alpha_K^*$   
**output:** updated  $\alpha$ , a transmission strategy  $\mathbf{r}$

- 1  $\alpha = \max\{\alpha, \sum_{j \in \mathcal{K}} \frac{L - |\mathcal{H}_j|}{K-1}, \sum_{j \in \mathcal{K}} 2L - |\mathcal{H}_j| - |\mathcal{H}_{\mathcal{K} \setminus \{j\}}|\};$
- 2 initiate a  $K$ -partition  $\mathcal{W}_K = \{\{j\} : j \in \mathcal{K}\}$  and a  $K$ -dimension transmission strategy  $\mathbf{r} = (0, \dots, 0);$
- 3 **repeat**
- 4    $\mathcal{U} = \text{FindMergeCand}(\mathcal{W}_K, \alpha);$
- 5    $\mathbf{r} = \text{UpdateRates}(\mathbf{r}, \mathcal{U});$
- 6   update  $\mathcal{W}_K$  by merging all  $\mathcal{X} \in \mathcal{U};$
- 7   **if**  $\alpha > \sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$  **then**
- 8      $\alpha = \alpha + 1;$
- 9     **terminate** 'repeat' loop and go to step 2;
- 10   **endif**
- 11 **until**  $|\mathcal{W}_K| = 2$  or  $\mathcal{U} = \emptyset;$
- 12  $\mathbf{r} = \text{UpdateRates}(\mathbf{r}, \mathcal{W}_K);$
- 13  $\Delta r = \max\{\mathbf{r}_K - \alpha, 0\};$
- 14 choose  $j' \in \mathcal{Y}$  such that  $r_{j'} - (L - |\mathcal{H}_{\mathcal{K} \setminus \{j'\}}|) \geq \Delta r;$
- 15  $r_{j'} = r_{j'} - \Delta r;$

---

#### Algorithm 2: FindMergeCand (find merging candidate)

---

**input :** a partition of client set  $\mathcal{W}_K$ , sum-rate  $\alpha$   
**output:**  $\mathcal{U}$ , a set contains all candidates for merge

- 1  $k = 1;$
- 2 **repeat**
- 3    $k = k + 1;$
- 4    $\mathcal{Z} = \{\mathcal{Y} : \xi_\alpha(\mathcal{Y}) < 0, \mathcal{Y} \text{ is a } k\text{-subset of } \mathcal{W}_K\};$
- 5   **if**  $\mathcal{Z} \neq \emptyset$  **then**
- 6      $\mathcal{U} = \mathcal{Y}, \text{ where } \xi_\alpha(\mathcal{Y}) \leq \xi_\alpha(\mathcal{Y}'), \forall \mathcal{Y}' \in \mathcal{Z};$
- 7   **endif**
- 8 **until**  $k = |\mathcal{W}_K| - 1$  or  $\mathcal{U}$  is assigned (i.e.,  $\mathcal{Z} = \emptyset$ );

---

where  $e = \arg \min\{v_\alpha(\mathcal{M} \cup \{u\}) - v_\alpha(\{u\}) : u \in \mathcal{K} \setminus \mathcal{M}\}$ . Let  $\mathcal{W}_K$  be a partition generated by the IM algorithm. For any  $\mathcal{X} \in \mathcal{W}_K$  such that  $\mathcal{X}$  is not a singleton, if we start the Queyranne's algorithm with  $\mathcal{M}^{(0)} = \mathcal{S} \subset \mathcal{X}$ , we will get  $\mathcal{M}^{(|\mathcal{X}| - |\mathcal{S}|)} = \mathcal{X}$ .<sup>9</sup> Due to the crossing submodularity of  $v_\alpha$  [5], at any iteration  $m \in \{2, \dots, K-1\}$  of Queyranne's algorithm [8]

$$v_\alpha(\mathcal{M}^{(m)}) + v_\alpha(\{j\}) \leq v_\alpha(\mathcal{M}^{(m)} \setminus \mathcal{S}) + v_\alpha(\mathcal{S} \cup \{j\}), \quad (10)$$

for all  $j \in \mathcal{K} \setminus \mathcal{M}^{(m)}$  and  $\mathcal{S}$  such that  $\emptyset \neq \mathcal{S} \subseteq \mathcal{M}^{(m-1)}$ .<sup>10</sup> Also, the clients in  $\mathcal{Y}$  merges only if  $\xi_\alpha(\mathcal{Y}) < 0$ .  $\mathcal{W}_K$  satisfies  $\sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X}) \leq \sum_{\mathcal{X} \in \mathcal{W}'_K} v_\alpha(\mathcal{X})$  for all other  $\mathcal{W}'_K$  such that  $|\mathcal{W}_K| = |\mathcal{W}'_K|$ . Alternatively speaking,  $\mathcal{W}_K$  generated by the IM algorithm incurs the minimum values of  $\sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$ . So, if  $\alpha \leq \sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$  holds for all  $\mathcal{W}_K$  in IM algorithm, it means universal recovery is achievable with sum-rate  $\alpha$ . Since  $\alpha$  is increased by 1 if condition  $\alpha \leq \sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$  is breached, the output must equal to  $\alpha_K^*$  if the the input  $\alpha \leq \alpha_K^*$ .<sup>11</sup>

Consider an NPS-CDE system having  $|\mathcal{U}|$  clients and has-sets  $\mathcal{H}_X, \forall \mathcal{X} \in \mathcal{U}$ . Based on (10), we have  $\mathcal{U} = \mathcal{W}_{\mathcal{U}}^*$ .

<sup>9</sup>See Appendix D for the proof and examples.

<sup>10</sup>See the examples in Appendix E.

<sup>11</sup>Also note that after step 1 in IM algorithm, condition (7) holds for all 2- and  $K$ -partitions of  $\mathcal{K}$ . Therefore, we do not need to check condition 7 for 2- and  $K$ -partitions in the rest steps of IM algorithm.

---

**Algorithm 3: UpdateRates (update rates)**


---

**input** : merge candidate set  $\mathcal{U}$ , transmission strategy  $\mathbf{r}$   
**output**: updated transmission strategy  $\mathbf{r}$

- 1  $\alpha_{\hat{\mathcal{Y}}} = \lceil \sum_{\mathcal{X} \in \mathcal{Y}} \frac{|\mathcal{H}_{\hat{\mathcal{Y}}}| - |\mathcal{H}_{\mathcal{X}}|}{|\mathcal{Y}| - 1} \rceil$ ;
- 2  $\Delta\alpha_{\hat{\mathcal{Y}}} = \sum_{\mathcal{X} \in \mathcal{Y}} (\alpha_{\hat{\mathcal{Y}}} - |\mathcal{H}_{\hat{\mathcal{Y}}}| - |\mathcal{H}_{\mathcal{X}}|) - \alpha_{\hat{\mathcal{Y}}}$ ;
- 3 choose  $\mathcal{X}' \in \mathcal{Y}$  such that the rates of clients in  $\mathcal{X}'$  have not been assigned;
- 4 **forall** the  $\mathcal{X} \in \mathcal{U}$  **do**
- 5      $R = \alpha_{\hat{\mathcal{Y}}} - |\mathcal{H}_{\hat{\mathcal{Y}}}| - |\mathcal{H}_{\mathcal{X}}|$ ;
- 6     **if**  $\mathcal{X} = \mathcal{X}'$  **then**  $R = R - \Delta\alpha_{\hat{\mathcal{Y}}}$   $\Delta r = \max\{R - r_{\mathcal{X}}, 0\}$ ;
- 7      $r_{j'} = r_{j'} + \Delta r$ , where  $j'$  is the client that is randomly chosen in set  $\mathcal{X}$ ;
- 8 **end**

---

According to (2), in UpdateRates algorithm,  $\Delta r$  for each  $\mathcal{X} \in \mathcal{U}$  determines the number of transmissions required from coalition  $\mathcal{X}$  for the local recovery in  $\mathcal{U}$  in addition to the local recovery in  $\mathcal{X}$ . The local recovery is achieved in every merged coalition in IM algorithm. Steps 13 to 15 in IM algorithm are to reduce the excessive rates  $\max\{R - r_{\mathcal{X}}, 0\}$  from the client  $j'$  such that the current rate  $r_{j'}$  is greater than  $L - |\mathcal{H}_{\mathcal{K} \setminus \{j'\}}|$ , the lower bound on the rate of client  $j'$  for universal recovery. Therefore, the output  $\mathbf{r}$  achieves universal recovery and has sum-rate equal to  $\alpha_{\mathcal{K}}^*$ . ■

## VI. COMPLEXITY

The complexity of the IM algorithm depends on two aspects. One is how close the input lower bound  $\alpha$  is to  $\alpha_{\mathcal{K}}^*$ , since the IM algorithm will be repeated for  $\alpha - \alpha_{\mathcal{K}}^* + 1$  times until it updates  $\alpha$  to  $\alpha_{\mathcal{K}}^*$ . The other is the complexity of FindMergeCand algorithm which may vary with different NPS-CDE systems. For example, if FindMergeCand returns  $\mathcal{U}$  containing 2-subsets,  $\beta$  is  $O(K^2 \cdot \gamma)$ . Here,  $\gamma$  is the complexity of running the cardinality function  $|\mathcal{H}_{\mathcal{X}}|$ . The authors in [5] also proposed a deterministic algorithm with complexity  $O(K \cdot \text{SFM}(K))$  for searching the the minimum sum-rate and minimum sum-rate strategy in CDE systems. Here,  $\text{SFM}(K)$  is the complexity of solving a submodular function minimization problem. To our knowledge, the algorithm proposed in [6] has the lowest complexity of  $\text{SFM}(K)$  is  $O(K^5 \cdot \gamma + K^6)$ . An experiment in [5] shows that the actual runtime by using MATLAB code is  $4 \cdot 10^{-3} \cdot K^{1.85}$  seconds on average.

We run an experiment to show the actual complexity of the IM algorithm. We set the number of packets  $L = 50$  and vary the number of clients  $K$  from 5 to 120. For each value of  $K$ , we repeat the procedure below for 100 times.

- randomly generate the has-sets  $\mathcal{H}_j$  for all  $j \in \mathcal{K}$  subject to the condition  $\cup_{j \in \mathcal{K}} \mathcal{H}_j = \mathcal{P}$ ;
- set  $\alpha$  to be the lower bound on  $\alpha_{\mathcal{K}}^*$  derived in [7]; run the IM algorithm in MATLAB.

In each repetition, we count the actual complexity in terms of  $\gamma$  and runtime (including the complexity of using algorithm in [7] to determine the lower bound on  $\alpha_{\mathcal{K}}^*$ ). We plot the average complexity over 100 repetitions in Fig. 3. It shows that the average complexity is about  $O(K^{3.15} \cdot \gamma)$ . The runtime of the IM algorithm is less than that of the deterministic algorithm in [5] when the number of clients is no greater than 94.

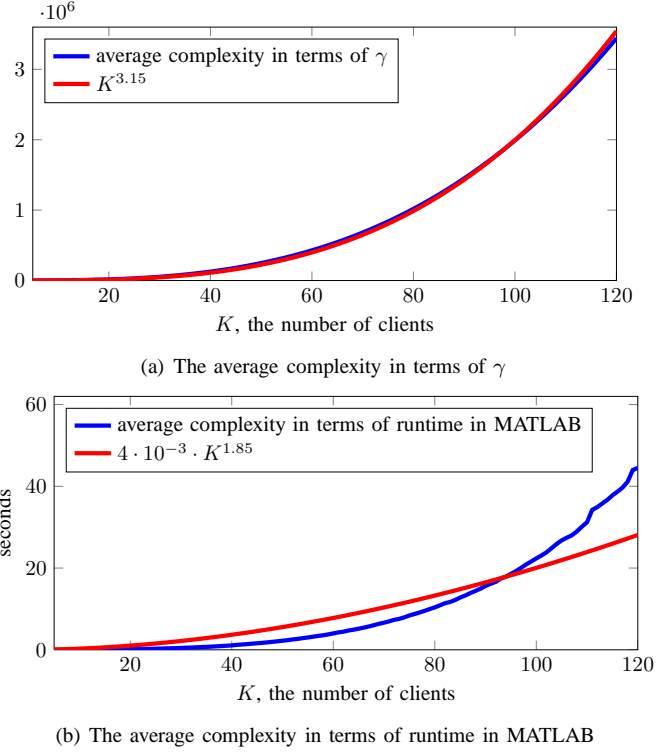


Fig. 3. The average complexity over 100 repetitions in experiment in Section VI.

## VII. CONCLUSION

This paper proposed an IM algorithm that found the minimum sum-rate and a minimum sum-rate strategy in NPS-CDE systems. The IM algorithm started with a sum-rate estimate, a lower bound on the minimum sum-rate. It recursively formed client sets into coalitions and updated the estimate to the value of minimum sum-rate. We proved that a minimum sum-rate strategy could be found by determine individual rate for achieving local recovery in each merged coalitions in the IM algorithm. Based on experiment results, we showed that the complexity of the IM algorithm was lower than the complexity of existing algorithms when the number of clients is below 94.

## REFERENCES

- [1] S. El Rouayheb, A. Sprintson, and P. Sadeghi, "On coding for cooperative data exchange," in *Proc. IEEE Inform. Theory Workshop*, Cairo, 2010, pp. 1–5.
- [2] A. Sprintson, P. Sadeghi, G. Booker, and S. El Rouayheb, "A randomized algorithm and performance bounds for coded cooperative data exchange," in *Proc. IEEE Int. Symp. Inform. Theory*, Austin, TX, 2010, pp. 1888–1892.
- [3] N. Abedini, M. Manjrekar, S. Shakkottai, and L. Jiang, "Harnessing multiple wireless interfaces for guaranteed QoS in proximate P2P networks," in *Proc. IEEE Int. Conf. Commun. China*, Beijing, 2012, pp. 18–23.
- [4] N. Milosavljevic, S. Pawar, S. El Rouayheb, M. Gastpar, and K. Ramchandran, "Deterministic algorithm for the cooperative data exchange problem," in *Proc. IEEE Int. Symp. Inform. Theory*, St. Petersburg, 2011, pp. 410–414.
- [5] T. Courtade and R. Wesel, "Coded cooperative data exchange in multihop networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1136–1158, Feb. 2014.
- [6] M. Goemans and V. Ramakrishnan, "Minimizing submodular functions over families of sets," *Combinatorica*, vol. 15, no. 4, pp. 499–513, 1995.
- [7] N. Ding, R. A. Kennedy, and P. Sadeghi, "Estimating minimum sum-rate for cooperative data exchange," *arXiv preprint arXiv:1502.03518*, 2015.

- [8] M. Queyranne, “Minimizing symmetric submodular functions,” *Math. Programming*, vol. 82, no. 1-2, pp. 3–12, Jun. 1998.

## Additional Pages (Appendices)

### APPENDIX A

Eq. (1) is based on the condition for the local recovery in  $\mathcal{S}$  [5]:

$$\mathbf{r}_{\mathcal{X}} \geq |\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{S} \setminus \mathcal{X}}|, \forall \mathcal{X} \subset \mathcal{S}, \quad (11)$$

which states that the total amount of information sent by set  $\mathcal{S} \setminus \mathcal{X}$  must be at least complementary to that collectively missing in set  $\mathcal{X}$ . Based on (11), we have

$$\alpha_{\mathcal{S}}^* \geq \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}} (|\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|), \forall \mathcal{W}_{\mathcal{S}}. \quad (12)$$

Likewise,  $\mathbf{r}_{\mathcal{S} \setminus \mathcal{X}} \geq |\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|$  should also be satisfied for all  $\mathcal{X} \subset \mathcal{S}$ . We have

$$\alpha_{\mathcal{S}}^* \geq \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}} \frac{|\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|}{|\mathcal{W}_{\mathcal{S}}| - 1}, \forall \mathcal{X} \subset \mathcal{S}. \quad (13)$$

It is shown in [7] that  $\sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}} \frac{|\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|}{|\mathcal{W}_{\mathcal{S}}| - 1} \geq \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{S}}} (|\mathcal{H}_{\mathcal{S}}| - |\mathcal{H}_{\mathcal{X}}|)$ , i.e., constraint (13) is tighter than (12). Also, because packet-splitting is not allowed in an NPS-CDE system,  $\alpha_{\mathcal{S}}^*$  must be an integer, we have (1).

### APPENDIX B

The divide-and-conquer (DV) algorithm iteratively divide non-singleton elements in the current partition by calling a 1-MAC function. The function 1-MAC( $\mathcal{S}$ ) returns two outputs:  $\alpha_{\mathcal{S}}^{\circ}$  and  $\mathcal{W}_{\mathcal{S}}^{\circ}$ , the maximum and maximizer of (3), respectively. The algorithm is shown in Algorithm 4.

In Algorithm 4,  $R_{\mathcal{X}} = \alpha_{\mathcal{S}}^{\circ} - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|$  is supposed to determine the sum-rate in  $\mathcal{X}$  for achieving local recovery in  $\mathcal{S}$ , which is based on Lemma 1 in [4] that the minimum sum-rate transmission strategy for the local recovery in  $\mathcal{S}$  satisfies  $\mathbf{r}_{\mathcal{X}} = \alpha_{\mathcal{S}}^{\circ} - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|$  for all  $\mathcal{X} \in \mathcal{W}_{\mathcal{S}}^{\circ}$ .  $\Delta r = R_{\mathcal{S}} - \alpha_{\mathcal{S}}^{\circ}$  calculates the difference between the sum-rate in  $\mathcal{S}$  that is required for achieving universal recovery in  $\mathcal{K}$  and the sum-rate in  $\mathcal{S}$  for achieving local recovery in  $\mathcal{S}$ . Based on Theorem 3 in [4],  $\Delta r \geq 0$  and it can be added to any set in  $\mathcal{W}_{\mathcal{S}}^{\circ}$ . The DV algorithm is repetitively called until the rate of each client is determined. It is claimed in [4] that  $\mathbf{r} = \text{DV}(\mathcal{K})$  is the minimum sum-rate strategy for an NPS-CDE system.<sup>12</sup> However, the counter example in Section IV shows that it is correct. In fact, one can show that  $\mathbf{r} = \text{DV}(\mathcal{K})$  is the minimum sum-rate strategy for a PS-CDE system instead of NPS-CDE system.

### APPENDIX C

$\mathcal{W}_{\mathcal{K}}^{\circ} = \{\{1\}, \{2\}, \{3\}, \{4\}\} = \{\{j\} : j \in \mathcal{K}\}$  implies: Among all constraints  $\mathbf{r}_{\mathcal{X}} \geq L - \mathcal{H}_{\mathcal{K} \setminus \mathcal{X}}$  for all  $\mathcal{X} \subset \mathcal{K}$  (for the universal recovery), the tightest constraints are  $\mathbf{r}_{\mathcal{K} \setminus \{j\}} \geq L - \mathcal{H}_{\{j\}}$  for all  $j \in \mathcal{K}$ .<sup>13</sup> Based on (1),  $\alpha_{\mathcal{K}}^*$  is an integer such that

$$\alpha_{\mathcal{K}}^* \geq \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{K}}} \frac{L - |\mathcal{H}_{\mathcal{X}}|}{|\mathcal{W}_{\mathcal{K}}| - 1} \quad (14)$$

<sup>12</sup>By calling  $\text{DV}(\mathcal{K})$ ,  $\mathcal{K}$  will be assigned to the first input argument  $\mathcal{S}$ .  $R_{\mathcal{S}}$  is not used if  $\mathcal{S} = \mathcal{K}$ .

<sup>13</sup>Alternatively speaking, a minimum sum-rate strategy must satisfy  $r_j \geq L - \mathcal{H}_{\mathcal{K} \setminus \{j\}}$  and  $\mathbf{r}_{\mathcal{K} \setminus \{j\}} \geq L - \mathcal{H}_{\{j\}}$  for all  $j \in \mathcal{K}$  and  $\mathbf{r}_{\mathcal{K}} = \alpha_{\mathcal{K}}^*$ .

**Algorithm 4: DV (divide-and-conquer)**


---

**input** :  $\mathcal{S}, R_{\mathcal{S}}$  if  $\mathcal{S} \neq \mathcal{K}$   
**output**:  $(R_j : j \in \mathcal{S})$ , where  $R_j$  determines transmission rate of client  $j$

- 1 initialize  $\Delta r = 0$ ;
- 2  $(\alpha_{\mathcal{S}}^{\circ}, \mathcal{W}_{\mathcal{S}}^{\circ}) = \text{1-MAC}(\mathcal{S})$ ;
- 3 **if**  $\mathcal{S} \neq \mathcal{K}$  **then**  $\Delta r = R_{\mathcal{S}} - \alpha_{\mathcal{S}}^{\circ}$  **forall the**  $\mathcal{X} \in \mathcal{W}_{\mathcal{S}}^{\circ}$  **do**  
 $R_{\mathcal{X}} = \alpha_{\mathcal{S}}^{\circ} - |\mathcal{H}_{\mathcal{S}}| + |\mathcal{H}_{\mathcal{X}}|$  randomly choose  $\mathcal{X}' \in \mathcal{W}_{\mathcal{S}}^{\circ}$ , do  
 $R_{\mathcal{X}'} = R_{\mathcal{X}'} + \Delta r$ ;
- 4 **forall the**  $\mathcal{X} \in \mathcal{W}_{\mathcal{S}}^{\circ}$  **such that**  $\mathcal{X}$  **is not singleton** **do**  
 $(R_j : j \in \mathcal{X}) = \text{DV}(\mathcal{X}, R_{\mathcal{X}})$ ;
- 5 **end**

---

for all  $\mathcal{W}_{\mathcal{K}}$  such that  $\mathcal{W}_{\mathcal{K}}$  is a partition of  $\mathcal{K}$  that satisfies  $2 \leq |\mathcal{W}_{\mathcal{K}}| \leq K$ . Equivalently,

$$\alpha_{\mathcal{K}}^* \leq \sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{K}}} v_{\alpha}(\mathcal{X}) \quad (15)$$

for all  $\mathcal{W}_{\mathcal{K}}$  such that  $2 \leq |\mathcal{W}_{\mathcal{K}}| \leq K$ , where  $v_{\alpha}(\mathcal{X}) = \alpha - L + |\mathcal{H}_{\mathcal{X}}|$ . In the NPS-CDE system in Example 4.1, we have

$$\sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{K}}} \frac{L - |\mathcal{H}_{\mathcal{X}}|}{|\mathcal{W}_{\mathcal{K}}| - 1} \leq \alpha_{\mathcal{K}}^* - \frac{1}{|\mathcal{W}_{\mathcal{K}}| - 1}$$

for all  $\mathcal{W}_{\mathcal{K}}$  such that  $|\mathcal{W}_{\mathcal{K}}| > 2$  and  $L - |\mathcal{H}_{\mathcal{X}}| + L - |\mathcal{H}_{\mathcal{K} \setminus \mathcal{X}}| = \alpha_{\mathcal{K}}^* - 1$  for all  $\mathcal{X} \in \mathcal{W}_{\mathcal{K}}^{\circ}$ . One can show that

$$v_{\alpha}(\mathcal{K} \setminus \mathcal{X}) = \min \left\{ \sum_{\mathcal{S} \in \mathcal{W}_{\mathcal{K} \setminus \mathcal{X}}} v_{\alpha}(\mathcal{S}) : \mathcal{W}_{\mathcal{K} \setminus \mathcal{X}} \text{ is a partition of } \mathcal{K} \setminus \mathcal{X} \right\} \quad (16)$$

for all  $\mathcal{X} \in \mathcal{W}_{\mathcal{K}}^{\circ} = \{\{1\}, \{2\}, \{3\}, \{4\}\}$ . For example, let  $\mathcal{X} = \{4\}$ . Based on  $\sum_{j \in \mathcal{K}} \frac{L - |\mathcal{H}_j|}{3} \leq \alpha_{\mathcal{K}}^* - \frac{1}{3}$  and  $L - |\mathcal{H}_{\{1,2,3\}}| + L - |\mathcal{H}_4| = \alpha_{\mathcal{K}}^* - 1$ , we have

$$L - |\mathcal{H}_1| + L - |\mathcal{H}_2| + L - |\mathcal{H}_3| \leq 3\alpha_{\mathcal{K}}^* - L + |\mathcal{H}_4| - 1 \\ = 2\alpha_{\mathcal{K}}^* + L - |\mathcal{H}_{\{1,2,3\}}|,$$

which is equivalent to  $v_{\alpha}(\{1,2,3\}) \leq v_{\alpha}(\{1\}) + v_{\alpha}(\{2\}) + v_{\alpha}(\{3\})$ . Based on  $\frac{L - |\mathcal{H}_{\{1,2\}}| + L - |\mathcal{H}_3| + L - |\mathcal{H}_4|}{2} \leq \alpha_{\mathcal{K}}^* - \frac{1}{2}$ , we can prove that  $v_{\alpha}(\{1,2,3\}) \leq v_{\alpha}(\{1,2\}) + v_{\alpha}(\{3\})$ . By this method, one can show that  $v_{\alpha}(\{1,2,3\}) \leq v_{\alpha}(\{1,3\}) + v_{\alpha}(\{2\})$  and  $v_{\alpha}(\{1,2,3\}) \leq v_{\alpha}(\{2,3\}) + v_{\alpha}(\{1\})$ . According to Proposition 3.1, we can choose any  $j \in \mathcal{K}$  to break  $\{1,2,3,4\}$  to  $\{\mathcal{K} \setminus \{j\}, \{j\}\}$  and consider how to achieve local recovery in  $\mathcal{K} \setminus \{j\}$ . For example, if we break client set to  $\{\{1,2,3\}, \{4\}\}$ , based on Proposition 3.1, we get  $r_{\{1,2,3\}} = 5$  and  $r_5 = 0$ . Then, consider the individual rates in  $\{1,2,3\}$  for the local recovery in  $\{1,2,3\}$ . We get  $r_1 = 3, r_2 = 2$  and  $r_3 = 0$ . One can show that  $(3, 2, 0, 0)$  is a minimum sum-rate strategy.

## APPENDIX D

THE PROOF AND EXAMPLES OF  $\mathcal{M}(|\mathcal{X}| - |\mathcal{S}|) = \mathcal{X}$  BY

STARTING QUEYRANNE'S ALGORITHM WITH

$\emptyset \neq \mathcal{M}^{(0)} = \mathcal{S} \subset \mathcal{X}$  IN THE PROOF OF THEOREM 5.3

Let  $\mathcal{S} \subset \mathcal{K}$  such that  $|\mathcal{S}| \leq K - 2$ . We have

$$v_{\alpha}(\mathcal{S} \cup \{u\}) - v_{\alpha}(\{u\}) = |\cup_{j \in \mathcal{S}} \mathcal{H}_j| - |\mathcal{H}_u \cap (\cup_{j \in \mathcal{S}} \mathcal{H}_j)|.$$

So,  $v_{\alpha}(\mathcal{S} \cup \{u\}) - v_{\alpha}(\{u\}) \leq v_{\alpha}(\mathcal{S} \cup \{u'\}) - v_{\alpha}(\{u'\})$  is equivalent to

$$|\mathcal{H}_u \cap (\cup_{j \in \mathcal{S}} \mathcal{H}_j)| \geq |\mathcal{H}_{u'} \cap (\cup_{j \in \mathcal{S}} \mathcal{H}_j)|.$$

Let  $\mathcal{W}_{\mathcal{K}}$  be the partition of  $\mathcal{K}$  that is generated by IM algorithm (at any iteration). For any  $\mathcal{X} \in \mathcal{W}_{\mathcal{K}}$ , since the clients  $u' \in \mathcal{K} \setminus \mathcal{X}$  is not merged to  $\mathcal{X}$ ,  $v_{\alpha}(\mathcal{S} \cup \{u\}) - v_{\alpha}(\{u\}) \leq v_{\alpha}(\mathcal{S} \cup \{u'\}) - v_{\alpha}(\{u'\})$ , for all  $\emptyset \neq \mathcal{S} \subset \mathcal{X}$ ,  $u \in \mathcal{X} \setminus \mathcal{S}$  and  $u' \in \mathcal{K} \setminus \mathcal{X}$ . For example, in Example 5.1, we have  $\mathcal{W}_{\mathcal{K}} = \{\{1,2,3\}, \{4\}\}$  in the second iteration. Consider  $\mathcal{X} = \{1,2,3\}$ . One can show that  $|\mathcal{H}_1 \cap \mathcal{H}_2| \geq |\mathcal{H}_2 \cap \mathcal{H}_4|$  and  $|(\mathcal{H}_1 \cup \mathcal{H}_2) \cap \mathcal{H}_3| \geq |(\mathcal{H}_1 \cup \mathcal{H}_2) \cap \mathcal{H}_4|$ .<sup>14</sup> They are equivalent to  $v_{\alpha}(\{2\} \cup \{1\}) \leq v_{\alpha}(\{2\} \cup \{4\})$  and  $v_{\alpha}(\{1,2\} \cup \{3\}) \leq v_{\alpha}(\{1,2\} \cup \{4\})$ , respectively.

Therefore, if  $\mathcal{M}^{(0)} = \mathcal{S}$ , we will get  $\mathcal{M}^{(|\mathcal{X}| - |\mathcal{S}|)} = \mathcal{X}$  at the  $|\mathcal{X}| - |\mathcal{S}|$ th iteration. For example, for  $\mathcal{W}_{\mathcal{K}} = \{\{1,2,3\}, \{4\}\}$  in Example 5.1 and  $\mathcal{X} = \{1,2,3\}$ , it can be shown that: If we start the Queyranne's algorithm with  $\mathcal{M}^{(0)} = \{1\}, \{2\}$  or  $\{3\}$ , we will get  $\mathcal{M}^{(2)} = \{1,2,3\}$ ; If we start the Queyranne's algorithm with  $\mathcal{M}^{(0)} = \{1,2\}, \{2,3\}$  or  $\{1,3\}$ , we will still get  $\mathcal{M}^{(1)} = \{1,2,3\}$ .

## APPENDIX E

EXAMPLES OF  $\mathcal{W}_{\mathcal{K}}$  GENERATED BY THE IM ALGORITHM INCURRING THE MINIMUM VALUES OF  $\sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{K}}} v_{\alpha}(\mathcal{X})$

*Example E.1:* Consider the NPS-CDE system in Fig. 1. We get  $\mathcal{W}_{\mathcal{K}} = \{\{1,2,3\}, \{4\}\}$  in the second iteration of IM algorithm. By applying Queyranne's algorithm with different  $\mathcal{M}^{(0)}$ , we have the following results:

- If  $\mathcal{M}^{(0)} = \{1\}$  or  $\mathcal{M}^{(0)} = \{3\}$ , then  $\mathcal{M}^{(1)} = \{1,3\}$  and  $\mathcal{M}^{(2)} = \{1,2,3\}$ . According to (10), we have

$$v_{\alpha}(\{1,2,3\}) + v_{\alpha}(\{4\}) \leq \begin{cases} v_{\alpha}(\{2\}) + v_{\alpha}(\{1,3,4\}) \\ v_{\alpha}(\{2,3\}) + v_{\alpha}(\{1,4\}) \\ v_{\alpha}(\{1,2\}) + v_{\alpha}(\{3,4\}) \end{cases} \quad (17)$$

- If  $\mathcal{M}^{(0)} = \{2\}$ , then  $\mathcal{M}^{(2)} = \{1,2,3\}$ . According to (10), we have

$$v_{\alpha}(\{1,2,3\}) + v_{\alpha}(\{4\}) \leq v_{\alpha}(\{1,3\}) + v_{\alpha}(\{2,4\}). \quad (18)$$

- If  $\mathcal{M}^{(0)} = \{1,2\}$ , then  $\mathcal{M}^{(1)} = \{1,2,3\}$ . According to (10), we have

$$v_{\alpha}(\{1,2,3\}) + v_{\alpha}(\{4\}) \leq v_{\alpha}(\{3\}) + v_{\alpha}(\{1,2,4\}). \quad (19)$$

- If  $\mathcal{M}^{(0)} = \{2,3\}$ , then  $\mathcal{M}^{(1)} = \{1,2,3\}$ . According to (10), we have

$$v_{\alpha}(\{1,2,3\}) + v_{\alpha}(\{4\}) \leq v_{\alpha}(\{1\}) + v_{\alpha}(\{2,3,4\}). \quad (20)$$

$v_{\alpha}(\{1,2,3\}) + v_{\alpha}(\{4\})$  and the right-hand-sides (RHSs) of Eqs. (17) to (20) are the values of  $\sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{K}}} v_{\alpha}(\mathcal{X})$  over all partitions  $\mathcal{W}_{\mathcal{K}}$  of  $\{1,2,3,4\}$  such that  $|\mathcal{W}_{\mathcal{K}}| = 2$ . Therefore,  $v_{\alpha}(\{1,2,3\}) + v_{\alpha}(\{4\})$  is the minimum value of  $\sum_{\mathcal{X} \in \mathcal{W}_{\mathcal{K}}} v_{\alpha}(\mathcal{X})$  among all 2-partitions.

<sup>14</sup>We show two examples of inequalities that can be derived based on  $\mathcal{W}_{\mathcal{K}} = \{\{1,2,3\}, \{4\}\}$  and  $\mathcal{X} = \{1,2,3\}$ . There are in fact many other such inequalities, e.g.,  $|\mathcal{H}_1 \cap \mathcal{H}_3| \geq |\mathcal{H}_3 \cap \mathcal{H}_4|$ ,  $|(\mathcal{H}_1 \cup \mathcal{H}_3) \cap \mathcal{H}_2| \geq |(\mathcal{H}_1 \cup \mathcal{H}_3) \cap \mathcal{H}_4|$ .



*Example E.2:* Consider an NPS-CDE system that contains 5 clients. They want to obtain a packet set that contains 10 packets. The has-sets are

$$\begin{aligned}\mathcal{H}_1 &= \{\mathbf{p}_5, \mathbf{p}_7, \mathbf{p}_{10}\}, \\ \mathcal{H}_2 &= \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_5, \mathbf{p}_6, \mathbf{p}_7, \mathbf{p}_8, \mathbf{p}_9\}, \\ \mathcal{H}_3 &= \{\mathbf{p}_1, \mathbf{p}_3, \mathbf{p}_5, \mathbf{p}_6, \mathbf{p}_7, \mathbf{p}_8, \mathbf{p}_9, \mathbf{p}_{10}\}, \\ \mathcal{H}_4 &= \{\mathbf{p}_1, \mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5, \mathbf{p}_6, \mathbf{p}_7, \mathbf{p}_8, \mathbf{p}_9\}, \\ \mathcal{H}_5 &= \{\mathbf{p}_3, \mathbf{p}_6, \mathbf{p}_8, \mathbf{p}_9\}.\end{aligned}$$

We get  $\mathcal{W}_K = \{\{3, 4\}, \{1\}, \{2\}, \{5\}\}$  in the first iteration and  $\mathcal{W}_K = \{\{2, 3, 4\}, \{1\}, \{5\}\}$  in the second iteration of the IM algorithm. Consider the partition  $\mathcal{W}_K = \{\{2, 3, 4\}, \{1\}, \{5\}\}$  in the second iteration. By applying Queyranne's algorithm with different  $\mathcal{M}^{(0)}$ , we have the following results:

- If  $\mathcal{M}^{(0)} = \{3\}$  or  $\mathcal{M}^{(0)} = \{4\}$ , then  $\mathcal{M}^{(1)} = \{3, 4\}$  and  $\mathcal{M}^{(2)} = \{2, 3, 4\}$ . According to (10), we have

$$\begin{aligned}v_\alpha(\{2, 3, 4\}) + v_\alpha(\{1\}) + v_\alpha(\{5\}) \\ \leq \begin{cases} v_\alpha(\{2\}) + v_\alpha(\{1, 3, 4\}) + v_\alpha(\{5\}) \\ v_\alpha(\{2\}) + v_\alpha(\{1\}) + v_\alpha(\{3, 4, 5\}) \\ v_\alpha(\{2, 3\}) + v_\alpha(\{1, 4\}) + v_\alpha(\{5\}) \\ v_\alpha(\{2, 3\}) + v_\alpha(\{1\}) + v_\alpha(\{4, 5\}) \\ v_\alpha(\{2, 4\}) + v_\alpha(\{1, 3\}) + v_\alpha(\{5\}) \\ v_\alpha(\{2, 4\}) + v_\alpha(\{1\}) + v_\alpha(\{3, 5\}) \end{cases} \quad (21)\end{aligned}$$

- If  $\mathcal{M}^{(0)} = \{2\}$ , then  $\mathcal{M}^{(2)} = \{2, 3, 4\}$ . According to (10), we have

$$\begin{aligned}v_\alpha(\{2, 3, 4\}) + v_\alpha(\{1\}) + v_\alpha(\{5\}) \\ \leq \begin{cases} v_\alpha(\{3, 4\}) + v_\alpha(\{1, 2\}) + v_\alpha(\{5\}) \\ v_\alpha(\{3, 4\}) + v_\alpha(\{1\}) + v_\alpha(\{2, 5\}) \end{cases} \quad (22)\end{aligned}$$

- If  $\mathcal{M}^{(0)} = \{2, 3\}$ , then  $\mathcal{M}^{(1)} = \{2, 3, 4\}$ . According to (10), we have

$$\begin{aligned}v_\alpha(\{2, 3, 4\}) + v_\alpha(\{1\}) + v_\alpha(\{5\}) \\ \leq \begin{cases} v_\alpha(\{4\}) + v_\alpha(\{1, 2, 3\}) + v_\alpha(\{5\}) \\ v_\alpha(\{4\}) + v_\alpha(\{1\}) + v_\alpha(\{2, 3, 5\}) \end{cases} \quad (23)\end{aligned}$$

- If  $\mathcal{M}^{(0)} = \{2, 4\}$ , then  $\mathcal{M}^{(1)} = \{2, 3, 4\}$ . According to (10), we have

$$\begin{aligned}v_\alpha(\{2, 3, 4\}) + v_\alpha(\{1\}) + v_\alpha(\{5\}) \\ \leq \begin{cases} v_\alpha(\{3\}) + v_\alpha(\{1, 2, 4\}) + v_\alpha(\{5\}) \\ v_\alpha(\{3\}) + v_\alpha(\{1\}) + v_\alpha(\{2, 4, 5\}) \end{cases} \quad (24)\end{aligned}$$

The RHSs of Eqs. (21) to (24) contain a set of minimum values of  $\sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$ . For example, consider the client subset  $\{1, 3, 4, 5\}$ . Since we  $\mathcal{W}_K = \{\{3, 4\}, \{1\}, \{2\}, \{5\}\}$  in the first iteration of the IM algorithm, one can show that  $v_\alpha(\{1, 3, 4\}) + v_\alpha(\{5\})$  and  $v_\alpha(\{1\}) + v_\alpha(\{3, 4, 5\})$  incur the minimum values of  $\sum_{\mathcal{X} \in \mathcal{W}_{\{1, 3, 4, 5\}}} v_\alpha(\mathcal{X})$  over all 2-partitions of  $\{1, 3, 4, 5\}$ . So,  $v_\alpha(\{2\}) + v_\alpha(\{1, 3, 4\}) + v_\alpha(\{5\})$  and  $v_\alpha(\{2\}) + v_\alpha(\{1\}) + v_\alpha(\{3, 4, 5\})$  at the RHS in (21) incur the minimum values of  $\sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$  over all 3-partitions of  $\mathcal{K} = \{1, 2, 3, 4, 5\}$  that contains a singleton  $\{2\}$ . Likewise,

one can show that  $v_\alpha(\{3\}) + v_\alpha(\{1, 2, 4\}) + v_\alpha(\{5\})$  and  $v_\alpha(\{3\}) + v_\alpha(\{1\}) + v_\alpha(\{2, 4, 5\})$  at the RHS in (24) incur the minimum values of  $\sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$  over all 3-partitions of  $\mathcal{K} = \{1, 2, 3, 4, 5\}$  that contains a singleton  $\{3\}$ .

Therefore,  $v_\alpha(\{2, 3, 4\}) + v_\alpha(\{1\}) + v_\alpha(\{5\})$  is the minimum value of  $\sum_{\mathcal{X} \in \mathcal{W}_K} v_\alpha(\mathcal{X})$  over all 3-partitions of  $\mathcal{K} = \{1, 2, 3, 4, 5\}$ .

## APPENDIX F

### ANOTHER EXAMPLE OF IM ALGORITHM

Consider a 4-client 8-packet NPS-CDE system with the has-sets being

$$\begin{aligned}\mathcal{H}_1 &= \{\mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_6, \mathbf{p}_7, \mathbf{p}_8\}, \mathcal{H}_2 = \{\mathbf{p}_1, \mathbf{p}_4, \mathbf{p}_7, \mathbf{p}_8\} \\ \mathcal{H}_3 &= \{\mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5, \mathbf{p}_6, \mathbf{p}_7, \mathbf{p}_8\}, \mathcal{H}_4 = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_6\}.\end{aligned}$$

We apply IM algorithm with input  $\alpha = 6$  and  $\alpha = 5$ . The procedures are shown in Examples F.1 and F.2, respectively.

*Example F.1:* Let  $\alpha = 6$ . We start IM algorithm. In step 1, we get  $\alpha = 6$ . Then, we initiate  $\mathcal{W}_K = \{\{1\}, \{2\}, \{3\}, \{4\}\}$  and  $\mathbf{r} = (0, 0, 0, 0)$ . Then, we have the following procedure.

- In the first iteration, we call FindNewPartition to determine  $\mathcal{U}$ . We first set  $k = 2$  and consider all 2-subsets.  $\{\{1\}, \{2\}\}$ ,  $\{\{1\}, \{3\}\}$  and  $\{\{2\}, \{3\}\}$  are the 2-subsets  $\mathcal{Y}$  that satisfy  $\xi_\alpha(\mathcal{Y}) < 0$ . Since

$$\xi_\alpha(\mathcal{Y}) = \begin{cases} -1 & \mathcal{Y} = \{\{1\}, \{2\}\} \\ -3 & \mathcal{Y} = \{\{1\}, \{3\}\}, \\ -1 & \mathcal{Y} = \{\{2\}, \{3\}\} \end{cases} \quad (25)$$

FindNewPartition returns  $\mathcal{U} = \{\{1\}, \{3\}\}$  to IM algorithm. We call UpdateRates algorithm, where  $r_1$  and  $r_3$  in  $\mathbf{r}$  is updated as  $r_1 = 0$  and  $r_3 = 1$ , respectively. So,  $\mathbf{r}$  is updated as  $\mathbf{r} = (0, 0, 1, 0)$ . We then merge sets  $\{1\}$  and  $\{3\}$  and update  $\mathcal{W}_K$  as  $\mathcal{W}_K = \{\{1, 3\}, \{2\}, \{4\}\}$ . Because  $\sum_{\mathcal{X} \in \mathcal{W}_K} v_6(\mathcal{X}) = v_6(\{1, 3\}) + v_6(\{2\}) + v_6(\{4\}) = 7 > \alpha$  for  $\mathcal{W}_K = \{\{1, 3\}, \{2\}, \{4\}\}$ ,  $\mathcal{U} \neq \emptyset$  and  $|\mathcal{W}_K| > 2$ , we continue the 'repeat' loop in IM algorithm.

- In the second iteration, FindNewPartition algorithm returns  $\mathcal{U} = \{\{1, 3\}, \{2\}\}$ . When calling UpdateRates algorithm, we have  $\Delta r = 2$  for  $\{1, 3\}$  and  $\Delta r = 1$  for  $\{2\}$ . we choose  $r_1$  to increase by two, and  $r_2$  is directly updated as  $r_2 = 1$ . The transmission strategy is updated as  $\mathbf{r} = (2, 1, 1, 0)$ , and  $\mathcal{W}_K$  is updated as  $\mathcal{W}_K = \{\{1, 2, 3\}, \{4\}\}$ . Since  $|\mathcal{W}_K| = 2$ , the 'repeat' loop in IM algorithm terminates. Note, in this iteration,  $\Delta r = 2$  for  $\{1, 3\}$  means that there should be 2 more transmission from client set, or coalition,  $\{1, 3\}$  for the local recovery in  $\{1, 2, 3\}$  in addition to the local recovery in  $\{1, 3\}$ . Also, since local recovery has been achieved in  $\{1, 3\}$ , client 1 and client 3 have the same has-set at this moment. Therefore,  $\Delta r = 2$  can be completed by any either of them. In this case we added  $\Delta r$  to client 1. But, it should be clear that local recovery in  $\{1, 2, 3\}$  can be achieved if we increase  $r_3$  by 2 or increase both  $r_1$  and  $r_3$  by 1. Since  $|\mathcal{W}_K| = 2$ , 'repeat' loop is terminated.
- We call UpdateRates by inputting  $\mathcal{U} = \mathcal{W}_K = \{\{1, 2, 3\}, \{4\}\}$ . We get  $\Delta r = 1$  for  $\{1, 2, 3\}$  and  $\Delta r = 1$



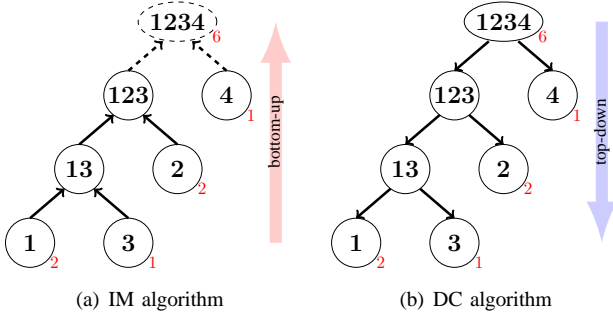


Fig. 4. The merging process results from iterative merging (IM) algorithm and dividing process results from divide-and-conquer (DV) algorithm when they are applied to find the minimum sum-rate strategy in the CDE system in Fig. 1. Note, final merging to one coalition  $\{1, 2, 3, 4\}$  does not happen but is implied in the IM algorithm. In each figure, the minimum sum-rate  $\alpha_K^*$  is shown beside the coalition  $\{1, 2, 3, 4\}$ , and the rates of clients in minimum sum-rate strategy are shown beside singleton coalitions. Note, the strategy determined by DC algorithm can not be implemented if packet-splitting is not allowed.

for  $\{4\}$ . We increase  $r_2$  by one and set  $r_4 = 1$ . The transmission strategy is updated as  $\mathbf{r} = (2, 2, 1, 1)$ .

Since  $\mathbf{r}_K = \alpha$ , the IM algorithm finally returns  $\alpha = 6$  and  $\mathbf{r} = (2, 2, 1, 1)$ . It can be shown that 6 is the minimum sum-rate and  $(2, 2, 1, 1)$  is one of the minimum sum-rate strategies.

*Example F.2:* Assume that we apply the IM algorithm to the CDE system in Fig. 1 with  $\alpha = 5$ . In step 1, we get  $\alpha = 5$ . However, it can be shown that  $\mathcal{W}_K = \{\{1, 3\}, \{2\}, \{4\}\}$  at the end of the first iteration and  $v_5(\{1, 3\}) + v_5(\{2\}) + v_5(\{4\}) = 4 < 5$ . 'repeat' loop terminates,  $\alpha$  is increased to 6 and the IM algorithm is started over again. With  $\alpha = 6$ , the same procedure as in Example F.1 is repeated.

We then apply the DV algorithm (in Algorithm 4) to this NPS-CDE system. The procedure and results are shown in the following example.

*Example F.3:* Consider the NPS-CDE system in Example F.1. The first call of  $1\text{-MAC}(\mathcal{K})$  returns  $\alpha_K^o = 6$  and  $\mathcal{W}_K^o = \{\{1, 2, 3\}, \{4\}\}$ , which gives transmission rates  $R_{\{1, 2, 3\}} = \alpha_K^o - L - |\mathcal{H}_{\{1, 2, 3\}}| = 5$  and  $R_{\{4\}} = \alpha_K^o - L - |\mathcal{H}_{\{4\}}| = 1$ . Then, the problem is to determine the exact rates of clients 1, 2 and 3. To do so,  $1\text{-MAC}(\{1, 2, 3\})$  is called, which returns  $\alpha_{\{1, 2, 3\}}^o = 4$  and  $\mathcal{W}_{\{1, 2, 3\}}^o = \{\{1, 3\}, \{2\}\}$ . We get  $R_{\{1, 3\}} = \alpha_{\{1, 2, 3\}}^o - |\mathcal{H}_{\{1, 2, 3\}}| + |\mathcal{H}_{\{1, 3\}}| = 3$  and  $R_{\{2\}} = \alpha_{\{1, 2, 3\}}^o - |\mathcal{H}_{\{1, 2, 3\}}| + |\mathcal{H}_{\{2\}}| = 1$  for the local recovery in  $\{1, 2, 3\}$ . In this case, there's an excessive rate  $\Delta r = R_{\{1, 2, 3\}} - \alpha_{\{1, 2, 3\}}^o = 1$  in set  $\{1, 2, 3\}$ .  $\Delta r$  will be added to any client in  $\{1, 2, 3\}$ , say, client 2, i.e.,  $r_2 = r_2 + 1 = 2$ . Then,  $1\text{-MAC}(\{1, 3\})$  is called. The results are  $\alpha_{\{1, 3\}}^o = 1$  and  $\mathcal{W}_{\{1, 3\}}^o = \{\{1\}, \{3\}\}$ , which means  $R_{\{1\}} = \alpha_{\{1, 3\}}^o - |\mathcal{H}_{\{1, 3\}}| + |\mathcal{H}_{\{1\}}| = 0$  and  $R_{\{3\}} = \alpha_{\{1, 3\}}^o - |\mathcal{H}_{\{1, 3\}}| + |\mathcal{H}_{\{3\}}| = 1$  are sufficient for the local recovery in  $\{1, 3\}$ . Let the excessive rate  $\Delta r = 2$  be added to client 1. We finally get the minimum sum-rate strategy  $\mathbf{r} = (2, 2, 1, 1)$ .

We show the merging and dividing processes resulted from the IM and the DV algorithms in Fig. 4. It can be shown that in this CDE system DV algorithm returns the same results as IM algorithm. And, in this case, the merging process of the IM algorithm is exactly the inverse procedure of the dividing process

of the DV algorithm. However, as explained in Section IV, DV algorithm is not applicable to CDE systems that do not allow packet-splitting in general.<sup>15</sup> From Fig. 4, we can see another problem with DV algorithm is that it is a top-down approach which could not be implemented in a decentralized manner. On the contrary, IM algorithm is a bottom-up approach, which allows clients to learn the transmission rates in a minimum sum-rate transmission strategy in a distributed/decentralized way.

<sup>15</sup>One can show that DV returns a minimum sum-rate strategy for CDE systems that do not allow packet-splitting if all  $\mathcal{W}_S^o$  returns by  $1\text{-MAC}(\mathcal{K})$  (in Algorithm 4) is a 2-partition. But, this is not necessarily the case in general.